UNIVERSITY OF ILLINOIS
AT URBANA-CHAMPAIGN

Basic Error Analysis

Physics 403

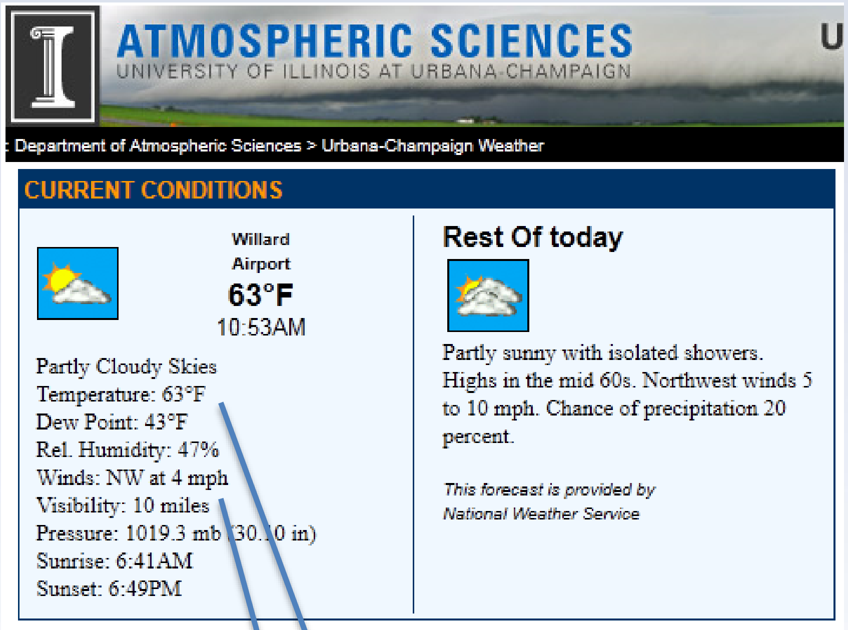
Summer 2016

Eugene V Colla & D Hertzog

Agenda

- Errors and uncertainties
- The Reading Error
- Accuracy and precession
- Systematic and statistical errors
- Fitting errors
- Rejection of the data points

What and when we need to know about errors. Everyday life.



$T = 63°\mathrm{F} \pm ?$ $\longrightarrow$ **Best guess $\Delta T \sim 0.5°\mathrm{F}$**

Wind speed 4mph$\pm$? $\longrightarrow$ **Best guess $\pm 0.5 mph$**

What and when we need to know about errors. Science.

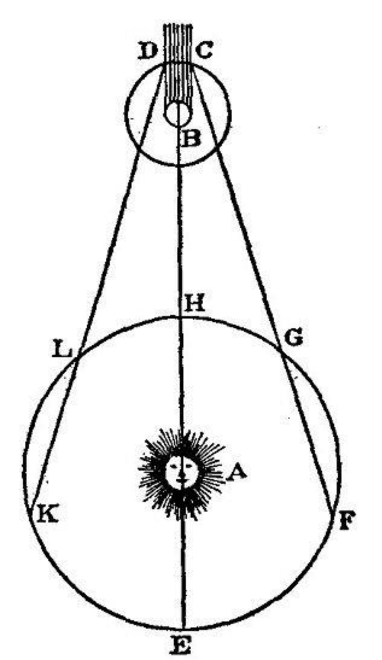


FIG. 70.

Measurement of the speed of the light

1675 Ole Roemer: 220,000 Km/sec

Ole Christensen Rømer
1644-1710

Does it make sense?
What is missing?

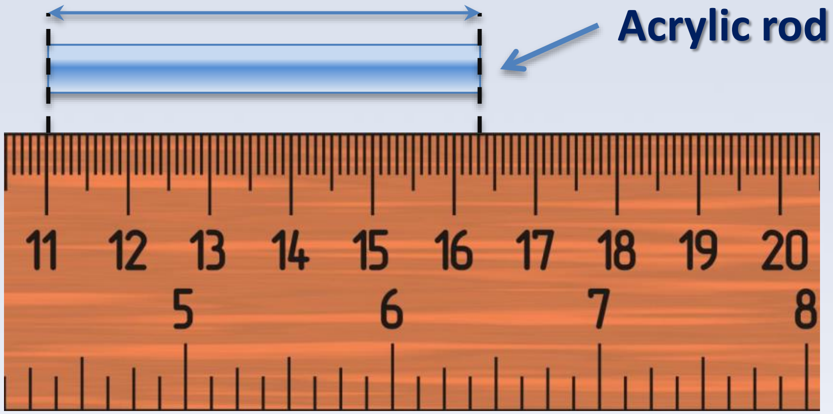
NIST Bolder Colorado c = 299,792,456.2±1.1 m/s.

Reading error

$\Delta L \cong 0.5mm$

L=53mm±ΔL(?)

Acrylic rod



$\Delta L \cong 0.03mm$

How far we have to go in reducing the reading error?

We do not care about accuracy better than 1mm

If ruler is not okay, we need to use digital caliper

Probably the natural limit of accuracy can be due to length uncertainty because of temperature expansion. For 53mm $\Delta L \cong 0.012mm/K$

Reading Error = ± $\frac{1}{2}$ (least count or minimum gradation).

Reading error. Digital meters.

Fluke 8845A multimeter

Example Vdc (reading)=0.85V

$$\Delta V = 0.83 \times (1.8 \times 10^{-5})$$
$$+1.0 \times (0.7 \times 10^{-5}) \cong 2.2 \times 10^{-5}$$
$$= 22\mu V$$

8846A Accuracy

Accuracy is given as ± (% measurement + % of range)

Range	24 Hour (23 ±1 °C)	90 Days (23 ±5 °C)	1 Year (23 ±5 °C)	Temperature Coefficient/ °C Outside 18 to 28 °C
100 mV	0.0025 + 0.003	0.0025 + 0.0035	0.0037 + 0.0035	0.0005 + 0.0005
1 V	0.0018 + 0.0006	0.0018 + 0.0007	0.0025 + 0.0007	0.0005 + 0.0001
10 V	0.0013 + 0.0004	0.0018 + 0.0005	0.0024 + 0.0005	0.0005 + 0.0001
100 V	0.0018 + 0.0006	0.0027 + 0.0006	0.0038 + 0.0006	0.0005 + 0.0001
1000 V	0.0018 + 0.0006	0.0031 + 0.001	0.0041 + 0.001	0.0005 + 0.0001

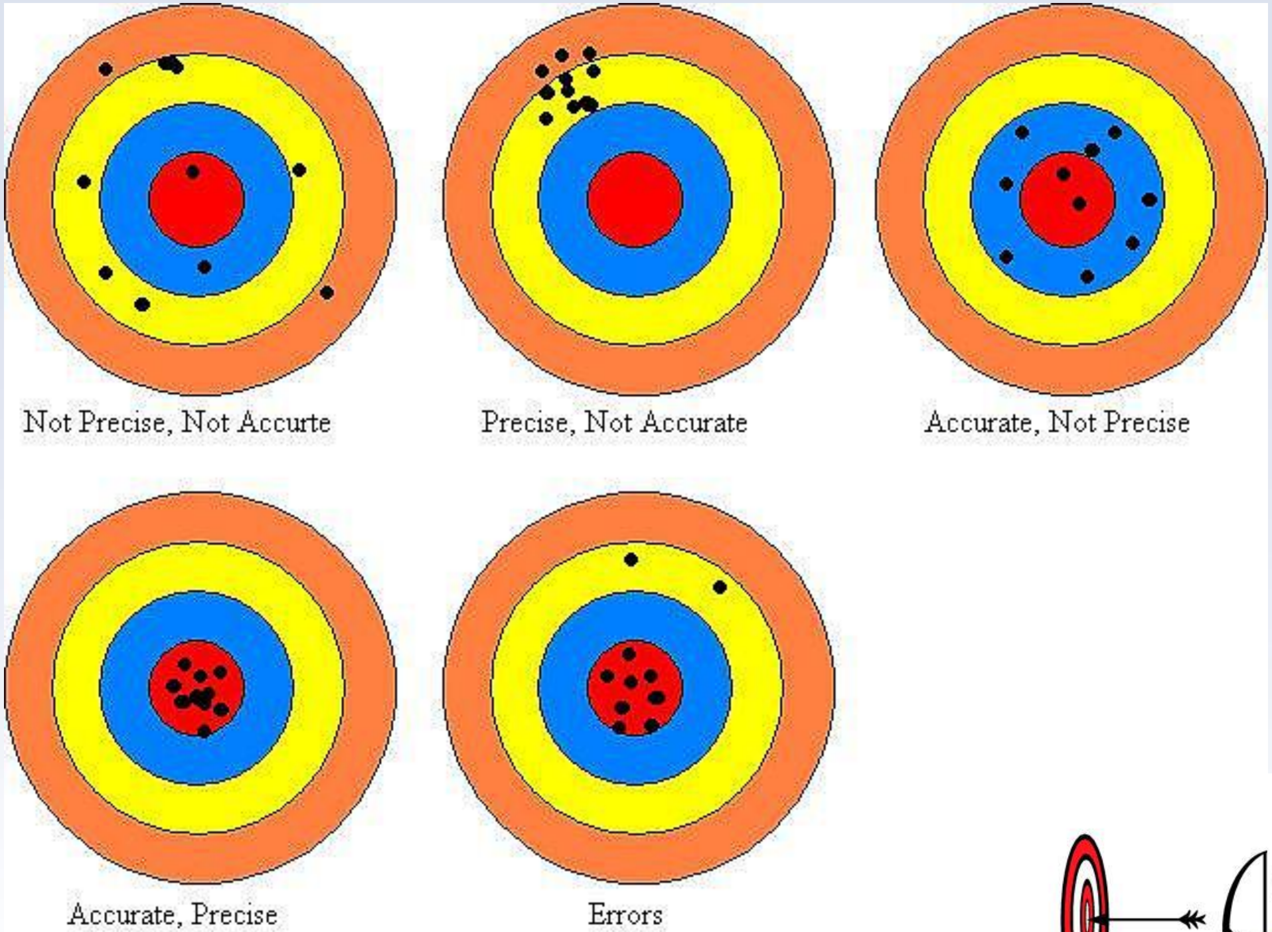
Accuracy and precision

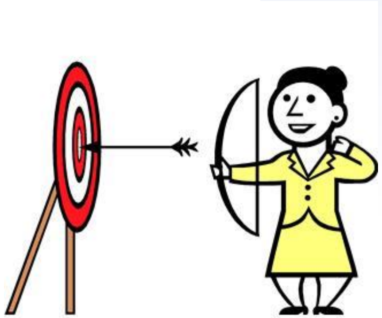
The accuracy of an experiment is a measure of how close the result of the experiment comes to the true value

Precision refers to how closely individual measurements agree with each other

illinois.edu

Accuracy and precision





Systematic and random errors

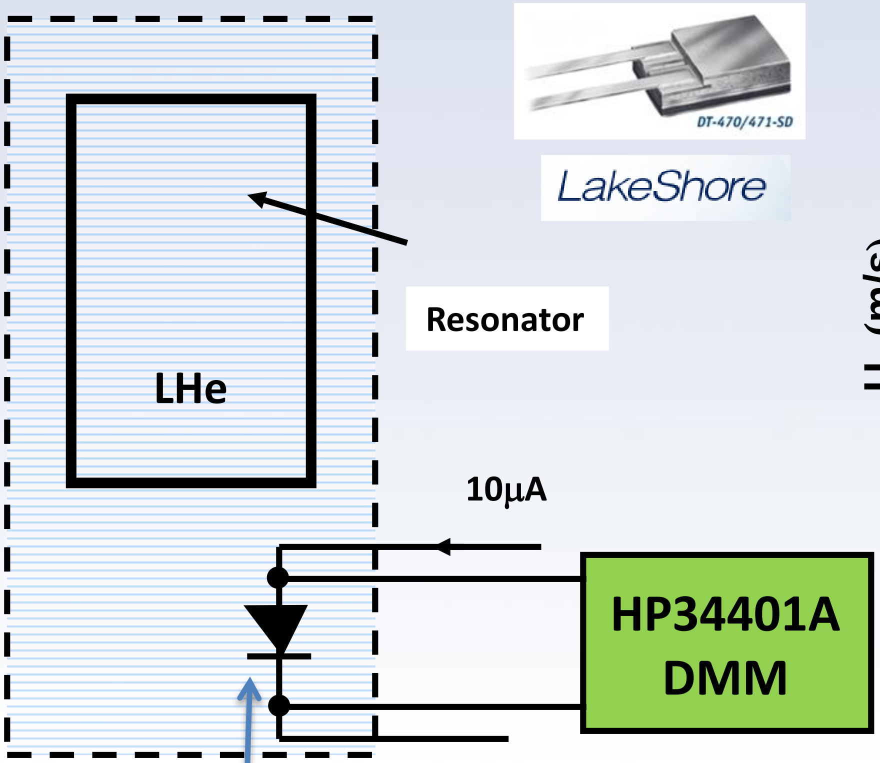
- ***Systematic Error:*** **reproducible inaccuracy introduced by faulty equipment, calibration or technique.**

- ***Random errors*: Indefiniteness of results due to finite precision of experiment. Measure of fluctuation in result after repeatable experimentation.**

Philip R. Bevington "Data Reduction and Error Analysis for the Physical sciences", McGraw-Hill, 1969

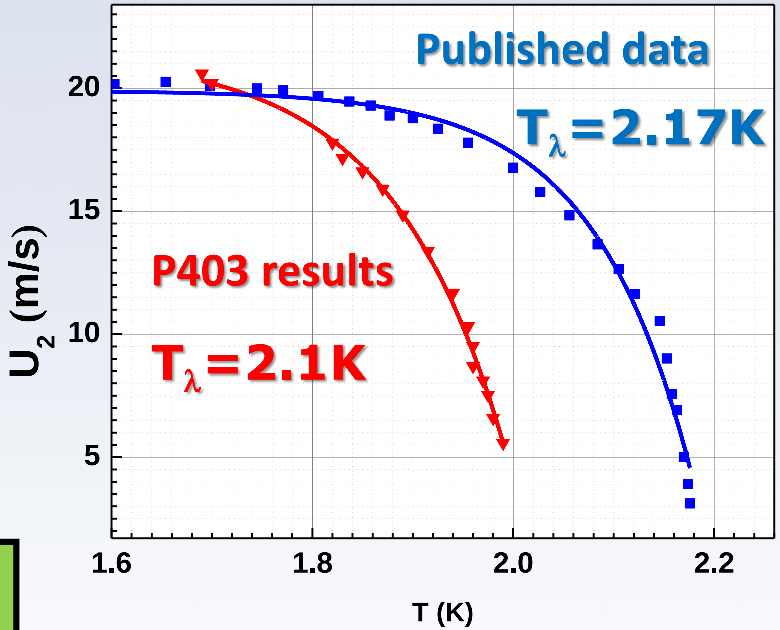
Systematic errors

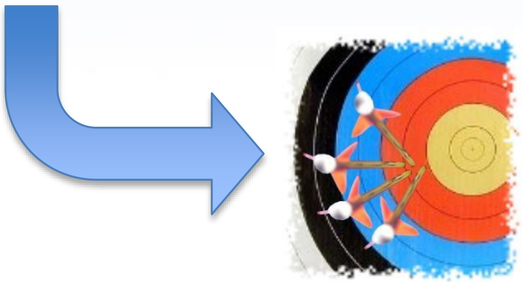
Example #3: poor calibration



Temperature sensor

Measuring of the speed of the second sound in superfluid He4





Random errors

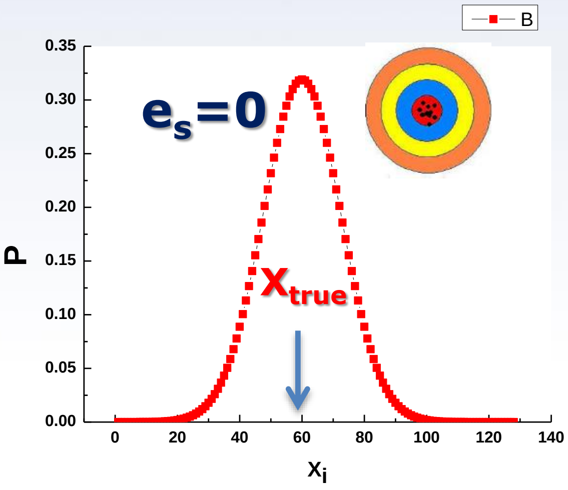
Result of measurement

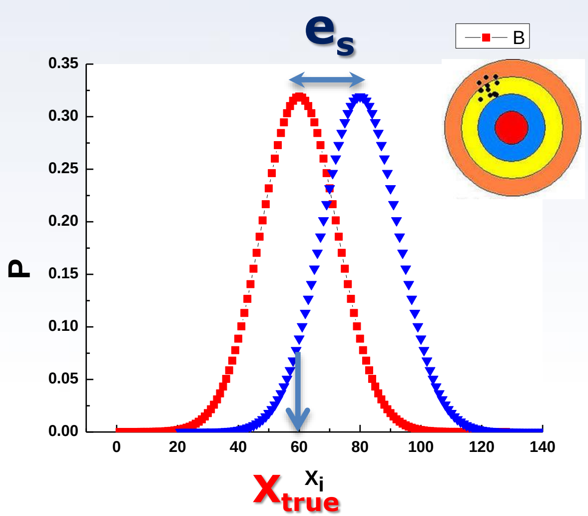
Systematic error

$$X_{meas} = X_{true} + e_s + e_r$$

Correct value

Random error





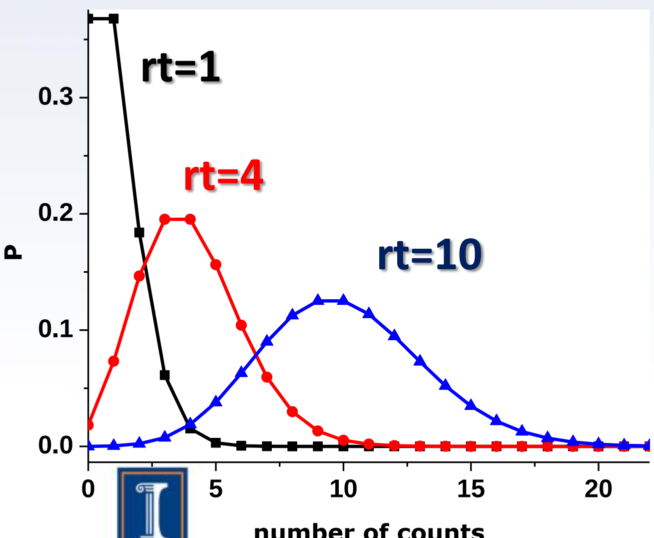
Random errors. Poisson distribution

Siméon Denis Poisson
(1781-1840)

$$P_n(t) = \frac{(rt)^n}{n!} e^{-rt} \qquad n = 0,1,2,\ldots$$

r: decay rate [counts/s] t: time interval [s]

➔ $P_n(rt)$: Probability to have n decays in time interval t



illinois.edu

A statistical process is described through a Poisson Distribution if:

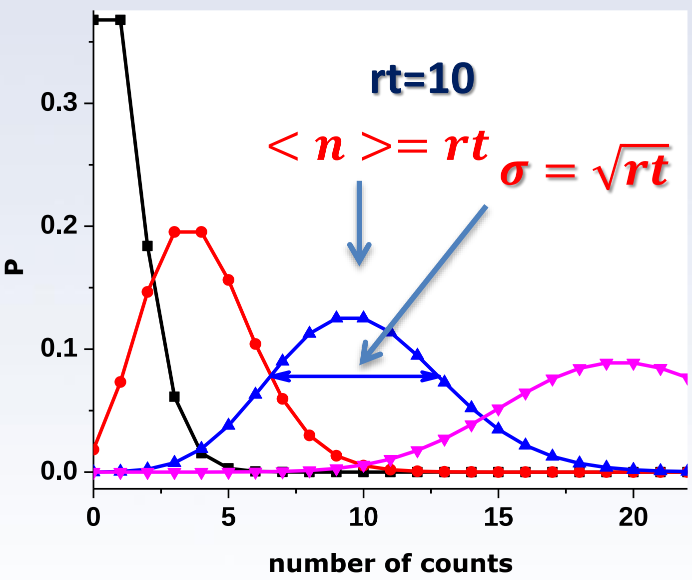
- **random process** → for a given nucleus probability for a decay to occur is the same in each time interval.
- **universal probability** → the probability to decay in a given time interval is same for all nuclei.
- **no correlation between two instances** (the decay of on nucleus does not change the probability for a second nucleus to decay.

Poisson distribution

$$P_n(t) = \frac{(rt)^n}{n!} e^{-rt} \qquad n = 0,1,2,\ldots$$

r: decay rate [counts/s] t: time interval [s]

➔ $P_n(rt)$: Probability to have n decays in time interval t



Properties of the Poisson distribution:

$$\sum_{n=0}^{\infty} P_n(rt) = 1$$, probabilities sum to 1

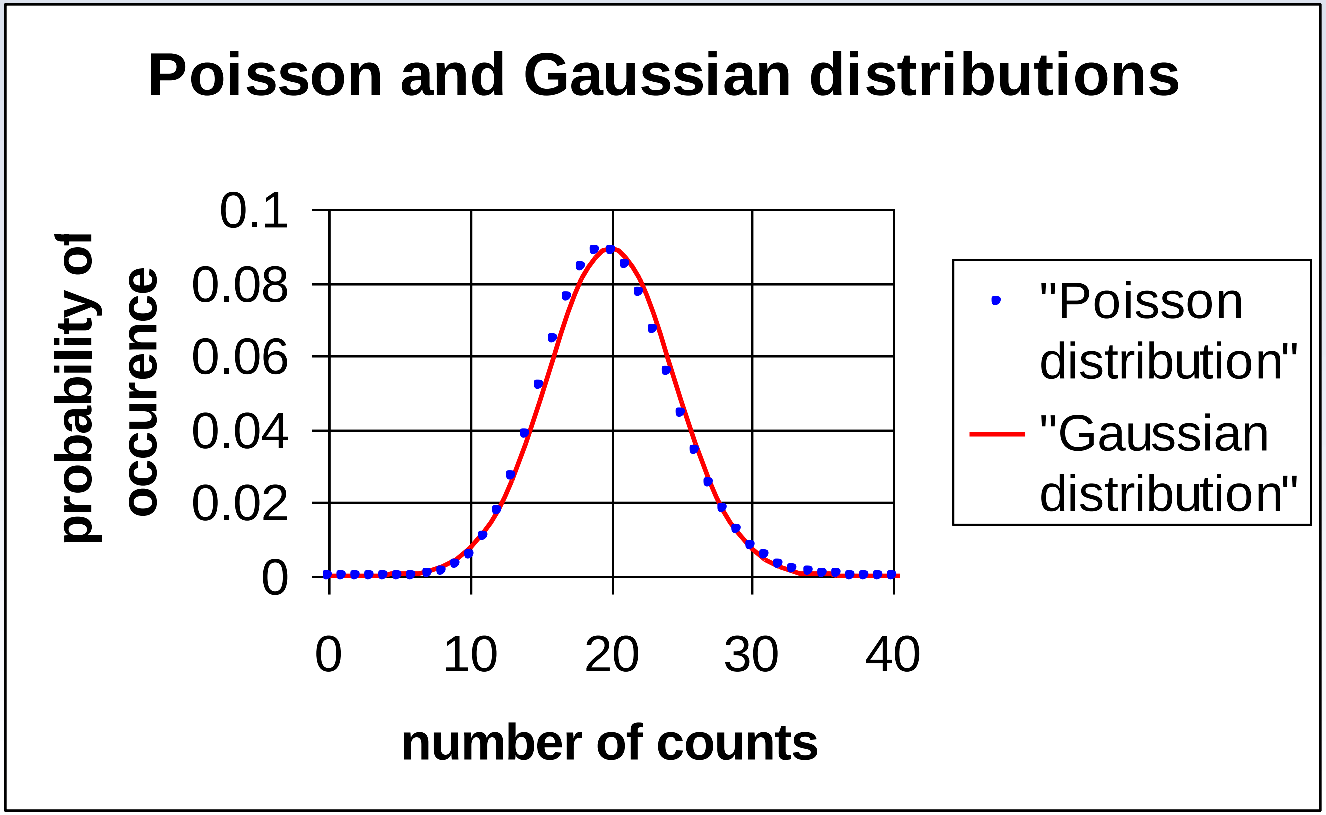
$$<n> = \sum_{n=0}^{\infty} n \cdot P_n(rt) = rt$$, the mean

$$\sigma = \sqrt{\sum_{n=0}^{\infty} (n - <n>)^2 P_n(rt)} = \sqrt{rt}$$, standard deviation

illinois.edu

Poisson distribution at large *rt*

$$P_n(t)=\frac{(rt)^n}{n!}e^{-rt}\quad n=0,1,2,\ldots$$

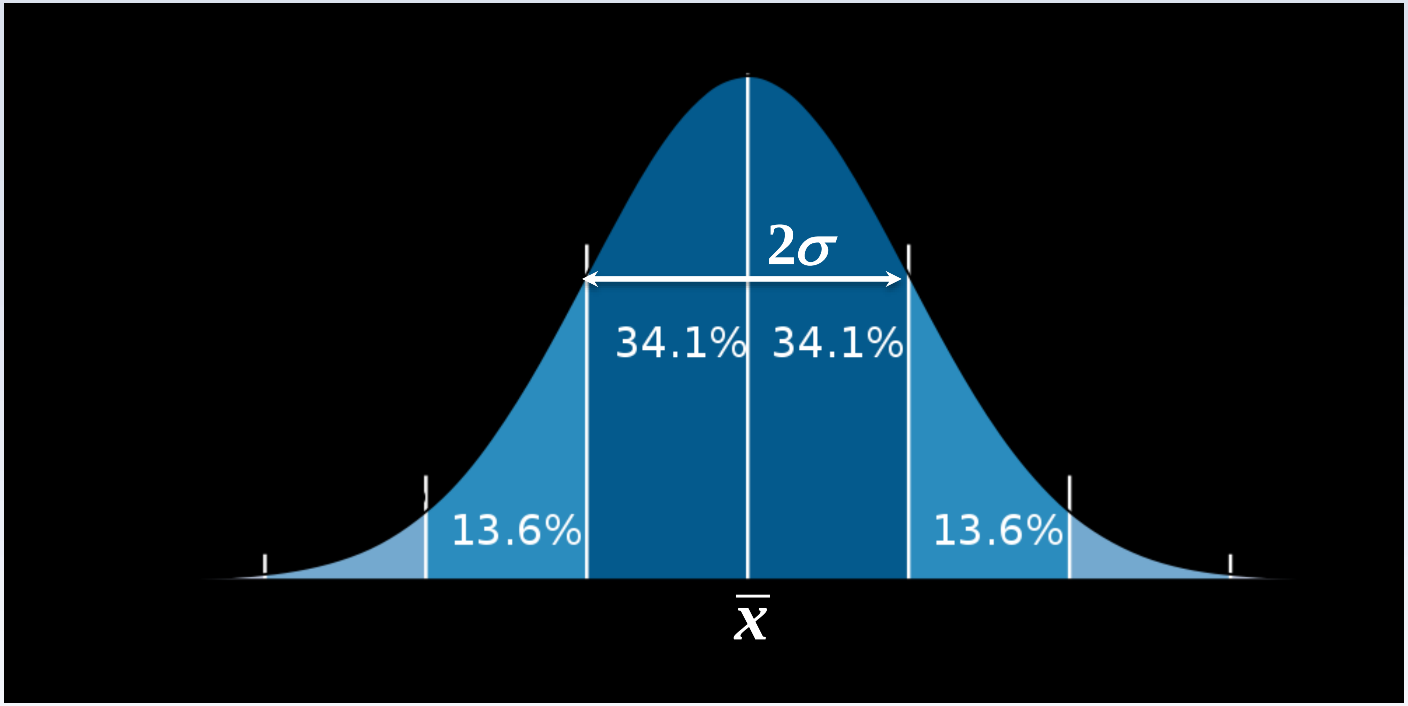


Carl Friedrich Gauss
(1777–1855)

$$P_n(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\bar{x})^2}{2\sigma^2}}$$

Gaussian distribution: continuous

Normal (Gaussian) distribution

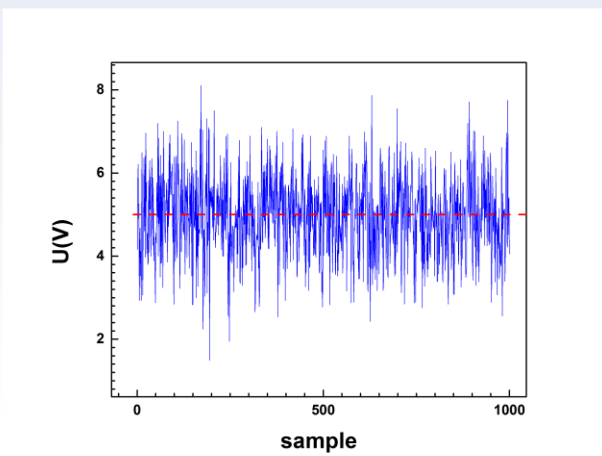


$$\boxed{P_n(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\bar{x})^2}{2\sigma^2}}}$$

Error in the mean is given as $\frac{\sigma}{\sqrt{N}}$

Measurement in presence of noise

Source of noisy signal

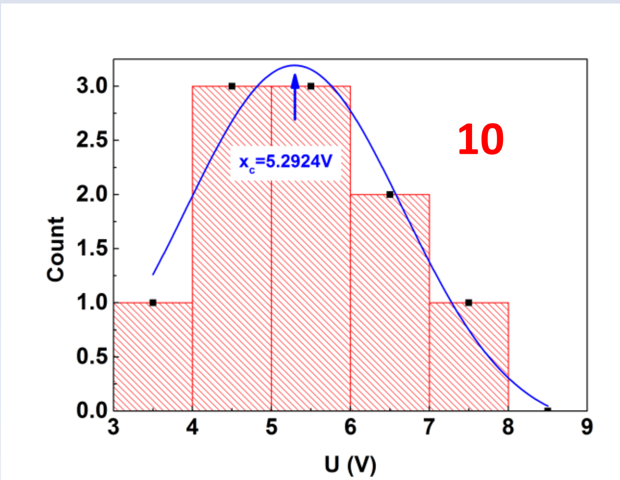


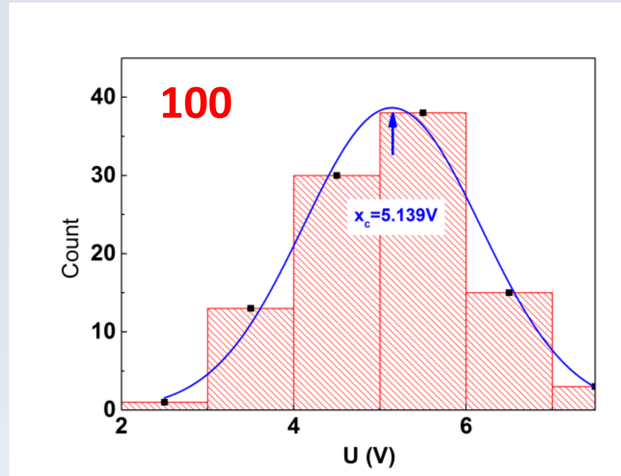
4.89855
5.25111
2.93382
4.31753
4.67903
3.52626
4.12001
2.93411

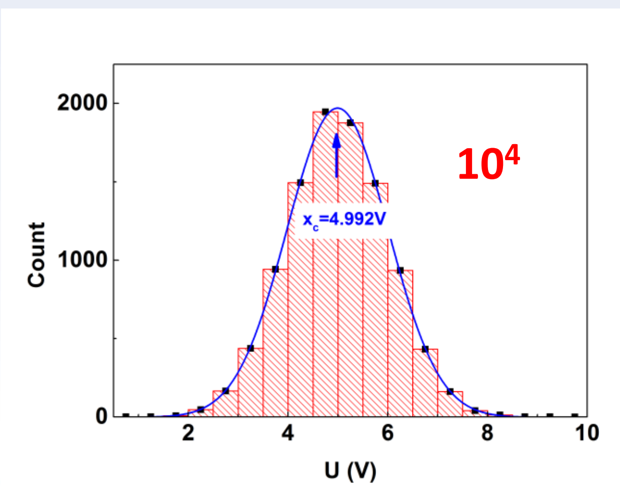
Expected value 5V

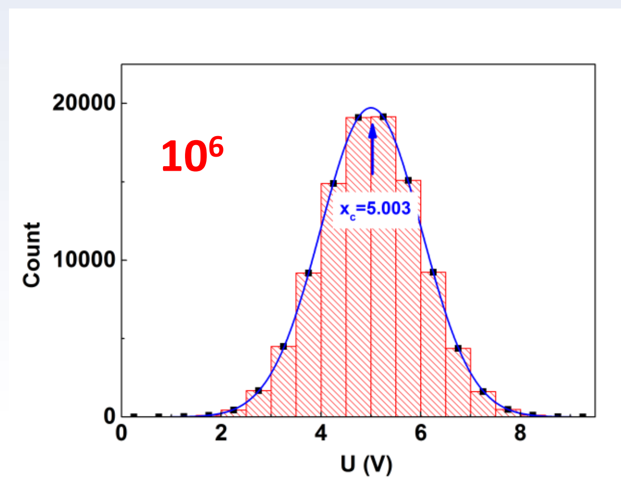
Actual measured values

Measurement in presence of noise



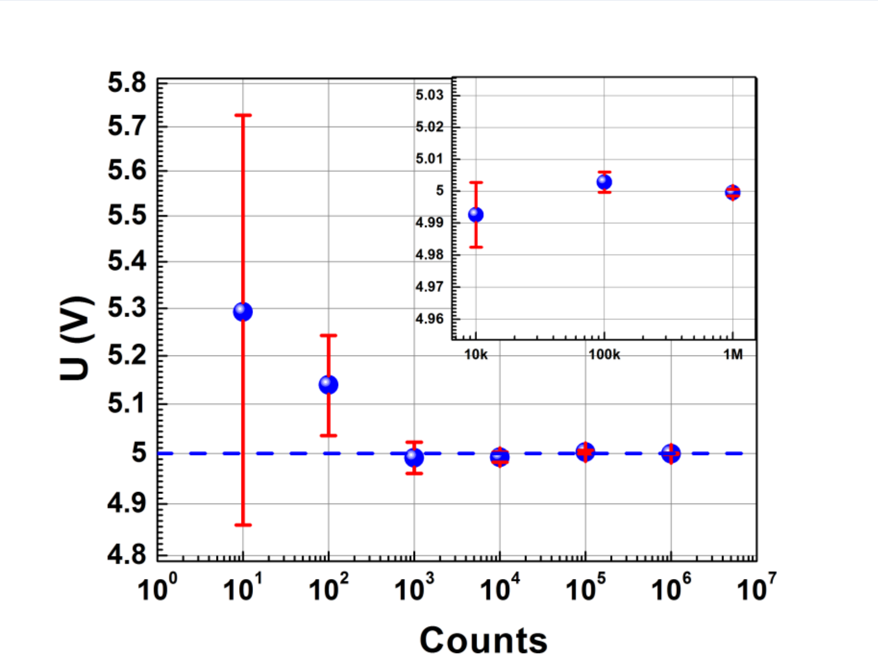


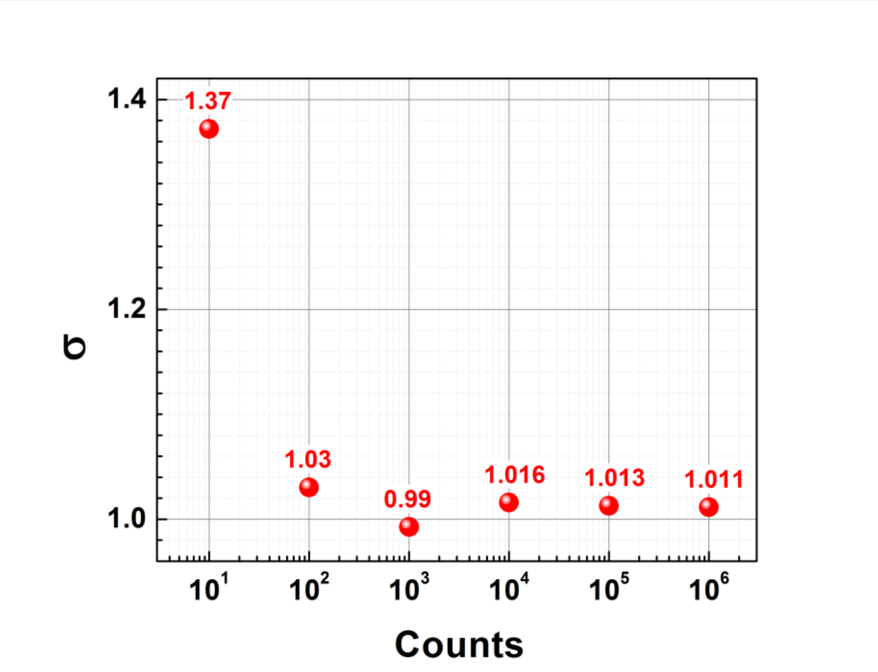




illinois.edu

Measurement in presence of noise





Result → $U = x_c \pm \frac{\sigma}{\sqrt{N}}$

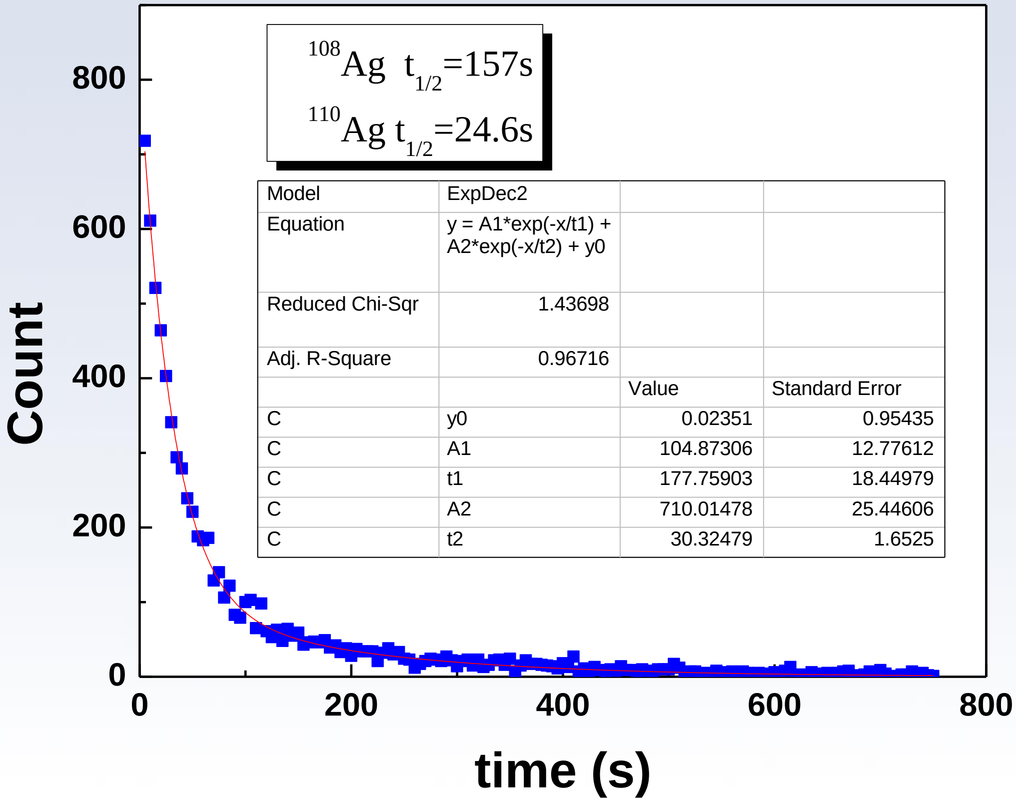
σ - standard deviation
N – number of samples

For N=10^6 U=4.999±0.001 0.02% accuracy

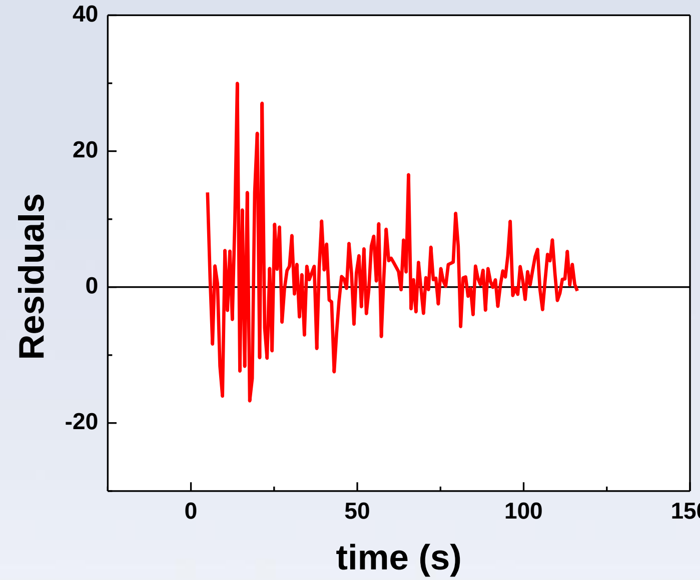
illinois.edu

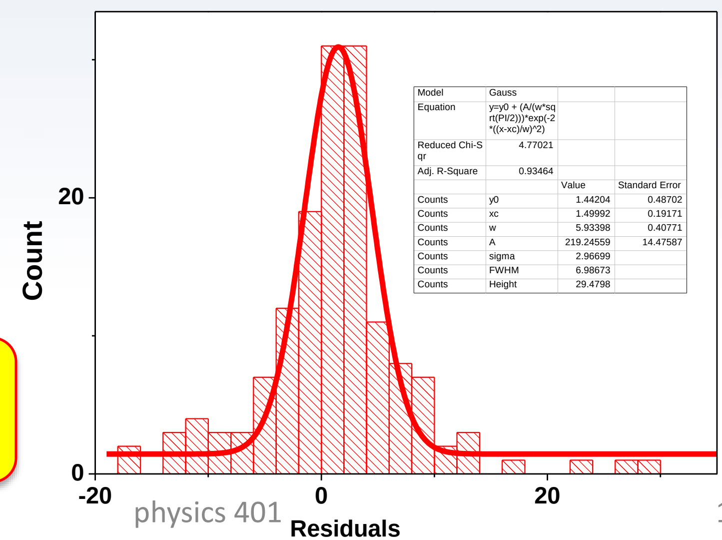
Fitting errors

Ag β decay



Model	ExpDec2		
Equation	y = A1*exp(-x/t1) + A2*exp(-x/t2) + y0		
Reduced Chi-Sqr	1.43698		
Adj. R-Square	0.96716		
		Value	Standard Error
C	y0	0.02351	0.95435
C	A1	104.87306	12.77612
C	t1	177.75903	18.44979
C	A2	710.01478	25.44606
C	t2	30.32479	1.6525





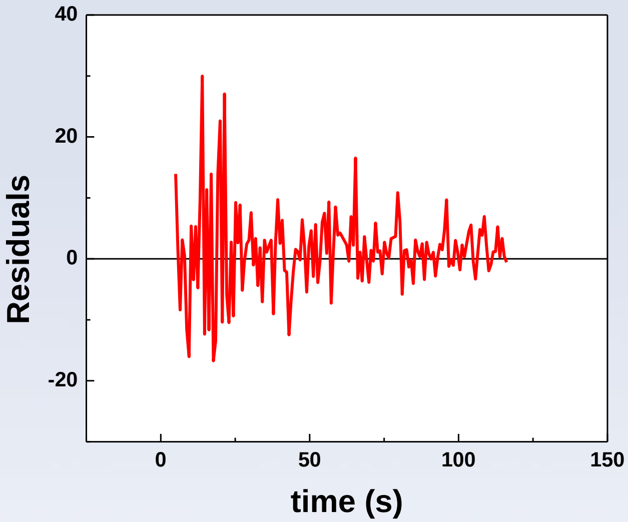
Model	Gauss		
Equation	y=y0 + (A/(w*sqrt(PI/2)))*exp(-2*((x-xc)/w)^2)		
Reduced Chi-Sqr	4.77021		
Adj. R-Square	0.93464		
		Value	Standard Error
Counts	y0	1.44204	0.48702
Counts	xc	1.49992	0.19171
Counts	w	5.93398	0.40771
Counts	A	219.24559	14.47587
Counts	sigma	2.96699	
Counts	FWHM	6.98673	
Counts	Height	29.4798	

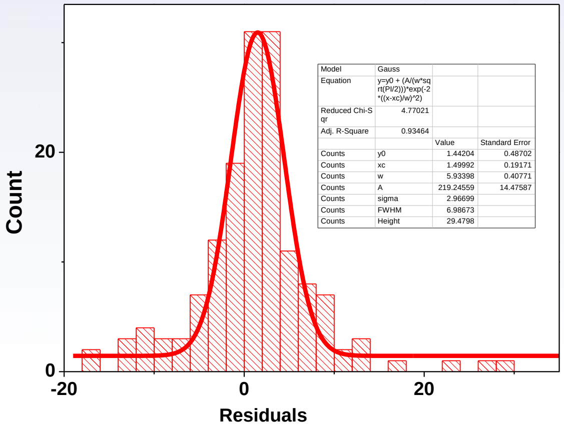
$$y = A1 \bullet \exp\left(\frac{-t}{t_1}\right) + A2 \bullet \exp\left(\frac{-t}{t_2}\right) + y_0$$

illinois.edu

Fitting. Analysis of the residuals

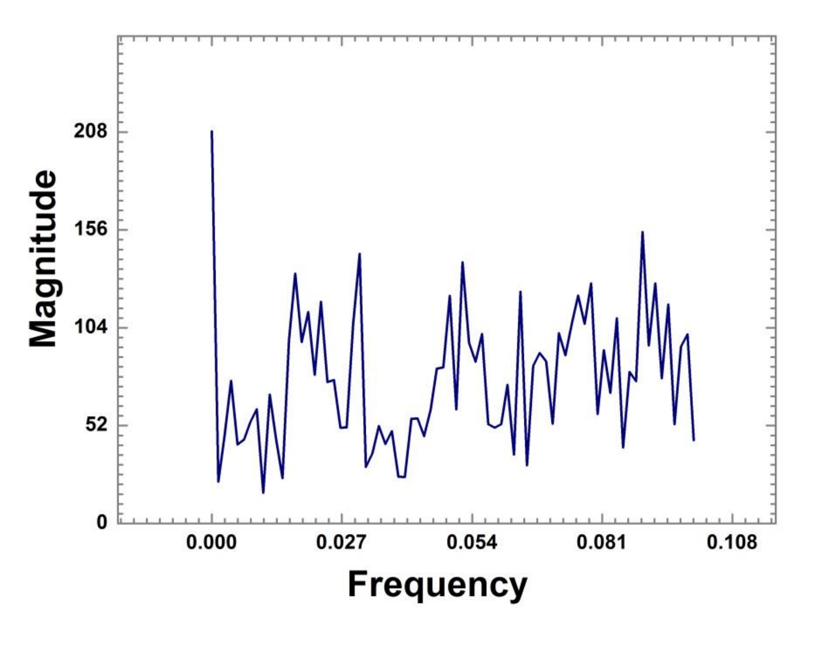
Ag β decay





Model	Gauss		
Equation	y=y0 + (A/(w*sqrt(PI/2)))*exp(-2*((x-xc)/w)^2)		
Reduced Chi-Sqr	4.77021		
Adj. R-Square	0.93464		
		Value	Standard Error
Counts	y0	1.44204	0.48702
Counts	xc	1.49992	0.19171
Counts	w	5.93398	0.40771
Counts	A	219.24559	14.47587
Counts	sigma	2.96699	
Counts	FWHM	6.98673	
Counts	Height	29.4798	

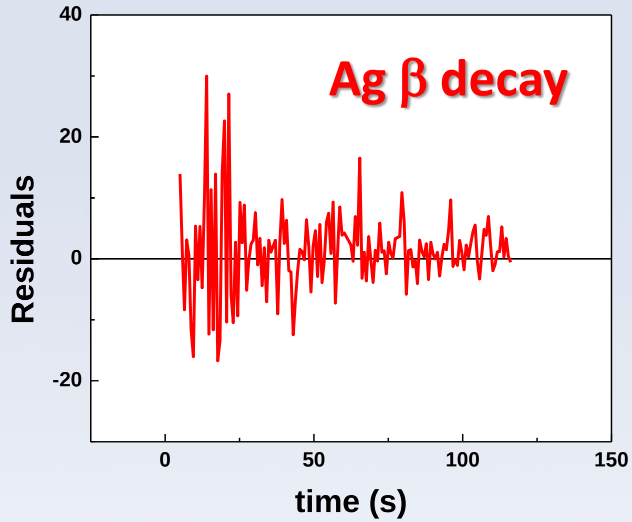
Test 1. Fourier analysis



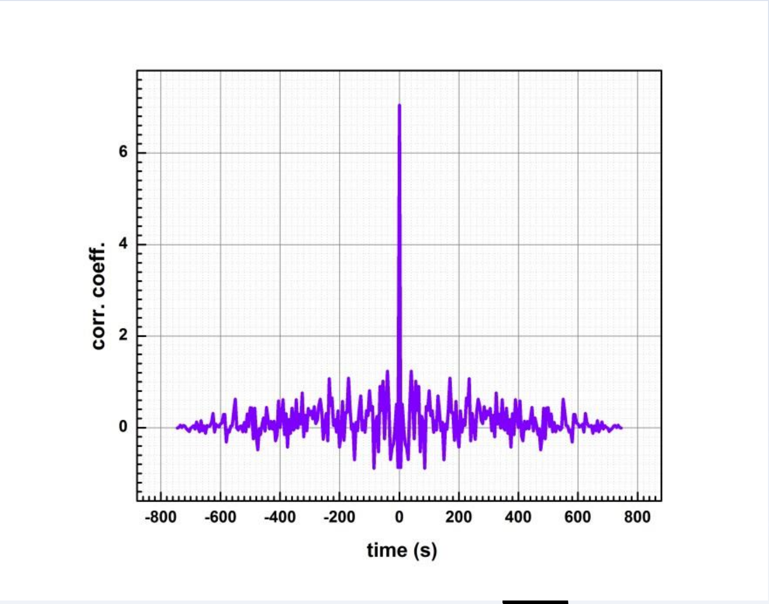
No pronounced frequencies found

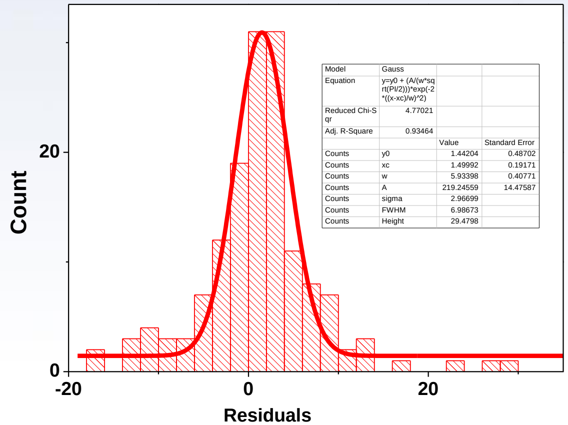
illinois.edu

Fitting. Analysis of the residuals



Test 1. Autocorrelation function





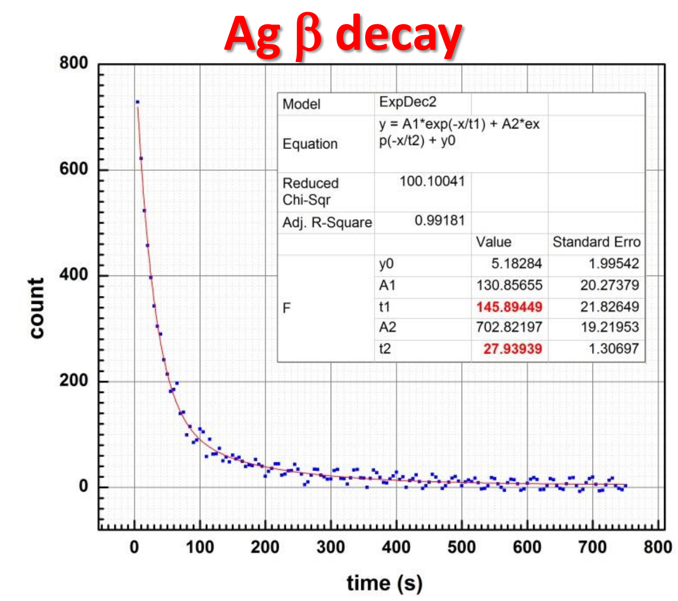
Correlation function

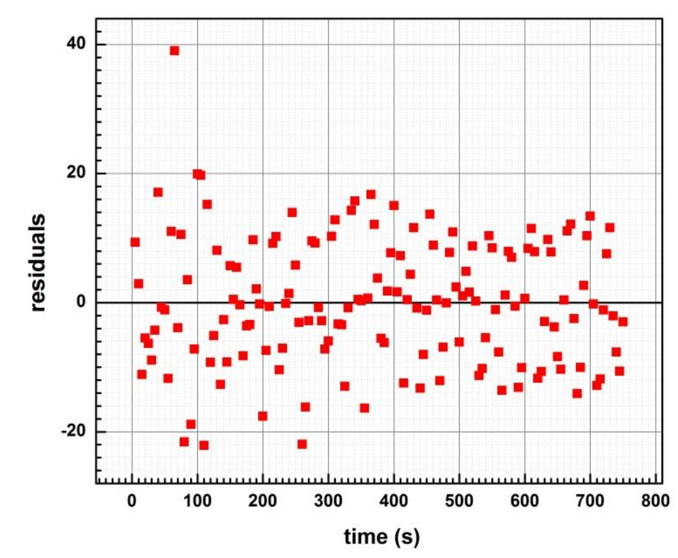
$$y(m) = \sum_{n=0} f(n) g(n-m)$$

autocorrelation function

$$y(m) = \sum_{n=0}^{M-1} f(n) f(n-m)$$

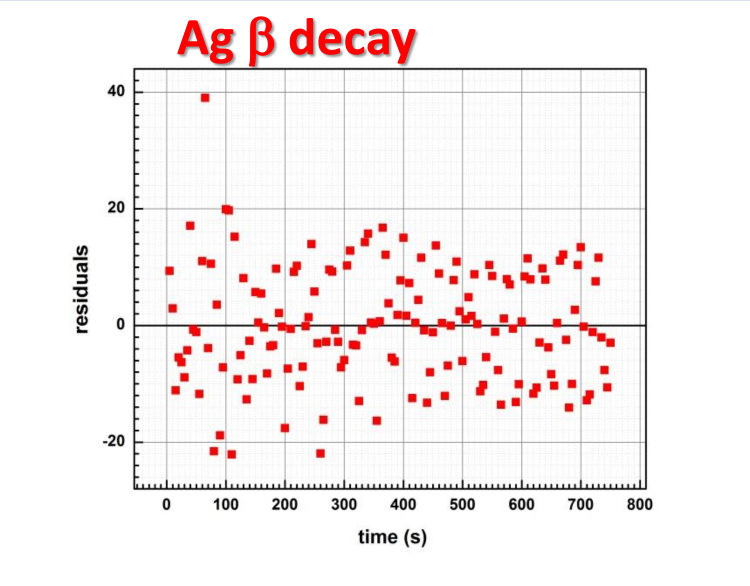
Fitting. Analysis of the residuals. Non “ideal” case

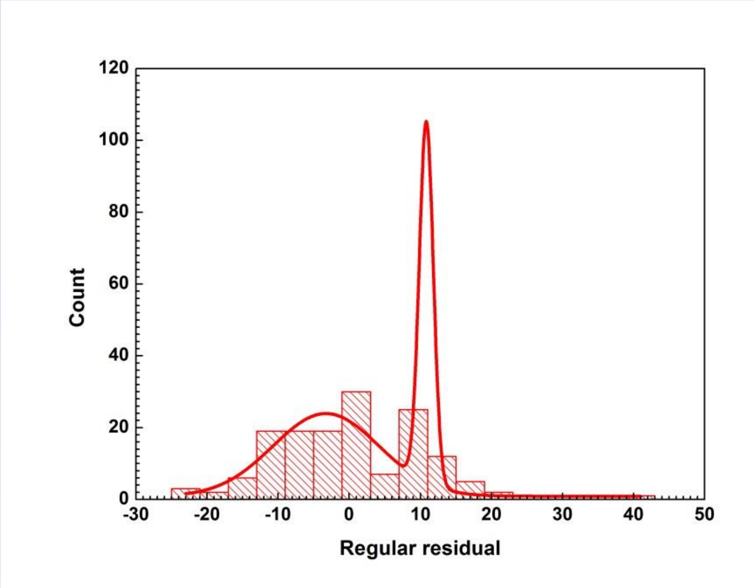


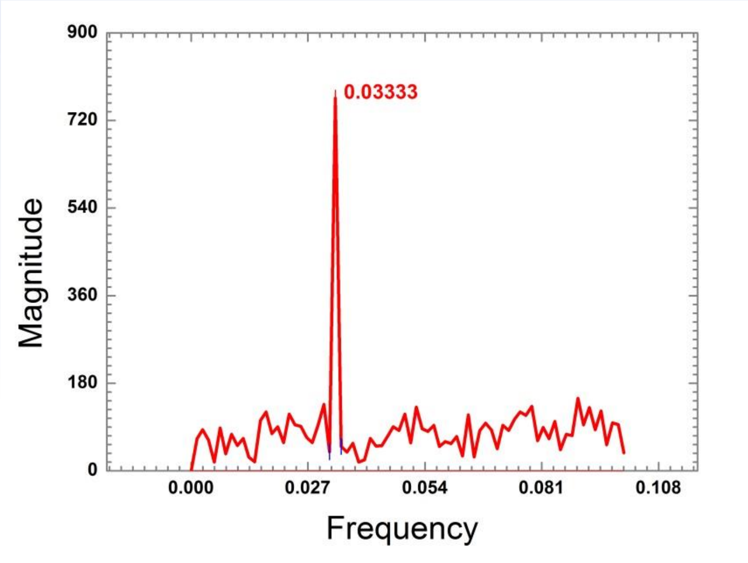


	Clear experiment	Data + “noise”
t_1(s)	177.76	145.89
t_2(s)	30.32	27.94

Fitting. Analysis of the residuals. Non “ideal” case

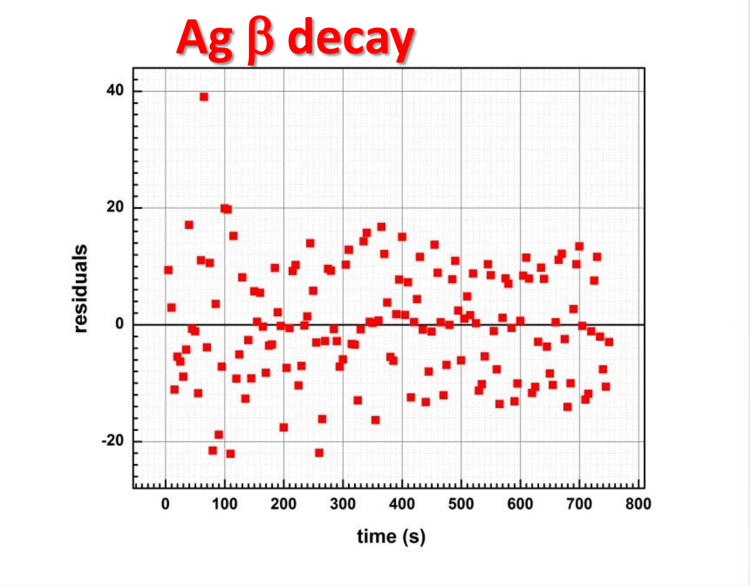


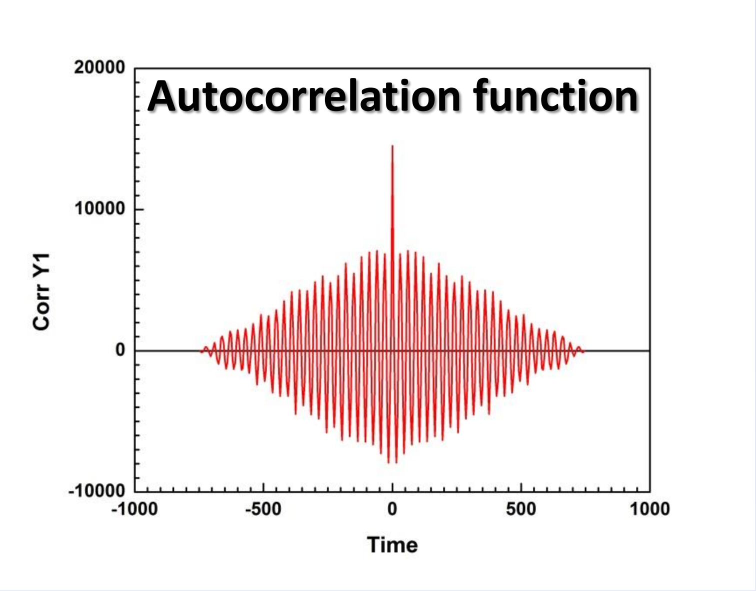




Histogram does not follow the normal distribution and there is frequency of 0.333 is present in spectrum

Fitting. Analysis of the residuals. Non "ideal" case

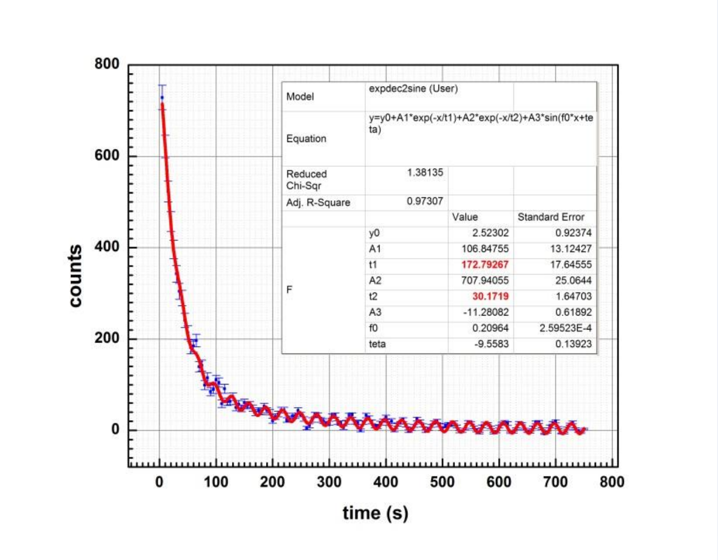


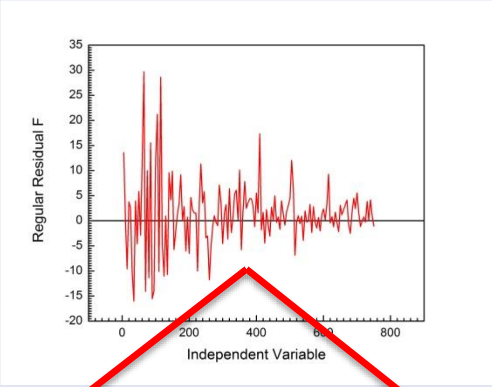


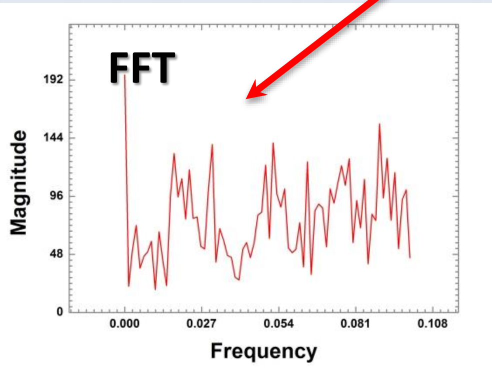
Conclusion: fitting function should be modified by adding an additional term:

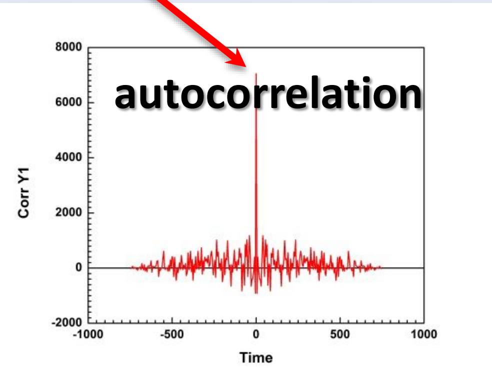
$$y(t) = y_0 + A_1 \exp\left(\frac{-t}{t_1}\right) + A_2 \exp\left(\frac{-t}{t_2}\right) + A_3 \sin(\omega t + \theta)$$

Fitting. Analysis of the residuals. Non "ideal" case







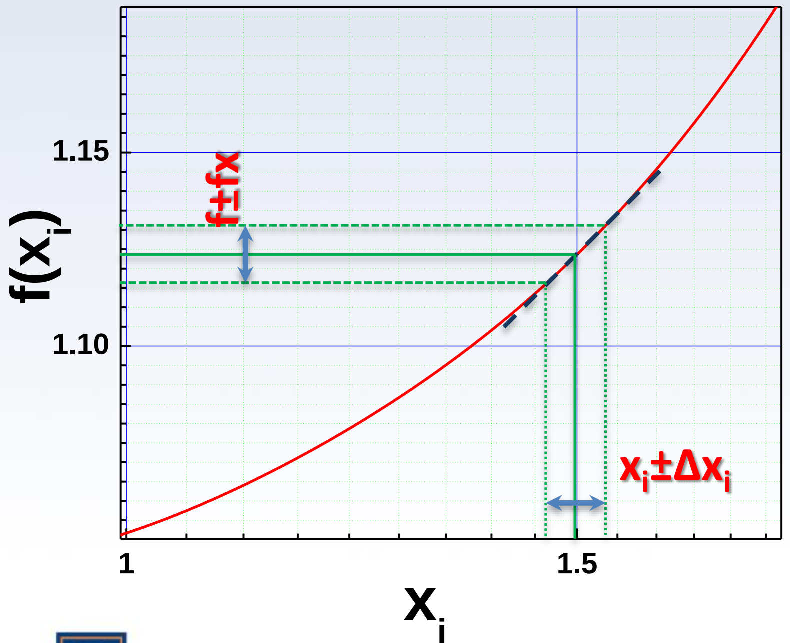


	Clear experiment	Data + noise	Modified fitting
t_1(s)	177.76	145.89	172.79
t_2(s)	30.32	27.94	30.17

Error propagation

y = f(x1, x2 ... xn)

$$\Delta f(x_i, \Delta x_i) = \sqrt{\sum_{i=1}^{n} \left[\frac{\partial f}{\partial x_i} \right]^2 \cdot \Delta x_i^2}$$



Error propagation. Example.

Derive resonance frequency *f* from measured inductance *L±ΔL* and capacitance *C±ΔC*

$$f(L,C) = \frac{1}{2\pi}\sqrt{\frac{1}{LC}}$$

$$L_1 = 10 \pm 1\text{mH}, \quad C_1 = 10 \pm 2\mu\text{F}$$

$$\Delta f(L,C,\Delta L,\Delta C) = \sqrt{\left[\frac{\partial f}{\partial L}\right]^2 \cdot \Delta L^2 + \left[\frac{\partial f}{\partial C}\right]^2 \Delta C^2}$$

$$\frac{\partial f}{\partial L} = \frac{-1}{4\pi} C^{-\frac{1}{2}} L^{-\frac{3}{2}};$$

$$\frac{\partial f}{\partial C} = \frac{-1}{4\pi} L^{-\frac{1}{2}} C^{-\frac{3}{2}}$$

Results:

$f(L_1,C_1)$=503.2921210448 7Hz

Δf=56.26977Hz

$f(L_1,C_1)$=503.3±56.3Hz

Rejection of Data

from J. Taylor, Ch. 6 of An Introduction to Error Analysis

Consider 6 measurements of a pendulum period : **3.8, 3.5, 3.9, 3.9, 3.4, 1.8**

Should the last measurement be rejected?

Yes: If some aspect of the experiment was changed ... new "slow" stopwatch, etc.

No: Never! You must always keep all data !! (diehards; beware)

Maybe? The usual case. You don't know why, but something may have made this measurement "bad." How do you set you set up a judgement that is unbiased?

First, compute some simple statistics:

Mean of measurements: $\bar{x} = 3.8$

Standard deviation: $$\sigma_x = \sqrt{\frac{1}{N}\sum(x_i - \bar{x})^2} = 0.8$$

Is the **1.8** measurement anomalous? It differs by 2σ (1.6) from the mean.

Chauvenet's Criterion

Chauvenet's criterion is the following: If the suspect measurement has a lower probability than 1/2, you should reject it.

We made 6 measurements. So, we expect that only 1/(20/6) = 0.3

New results:

Mean = **3.7**

Standard deviation = **0.2** ! much smaller !

Rejection of Data
Is all data good data?
No!

- Write down everything
  - in the logbook; take your time; use sentences; record numbers (values);
  - glitch in the power? note the time
  - temperature “cold” or “hot”? comment about it
  - somebody “reset” the system? note it please and when
- Record (electronically if possible) everything reasonable
  - as parallel information to the main data set
  - temperatures; voltages; generally called “slow controls”
- You *WILL* (almost certainly) have to go back and hunt for this documentation when something possibly anomalous arises … and it will